DOCUMENT RESUME

ED 430 220 CS 013 582

AUTHOR Slack, Jill Berlin; St. John, Edward P.
TITLE A Practical Model for Measuring the Effect of School Reform on the Reading Achievement of Non-Transient Learners.
PUB DATE 1999-04-22
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Canada, April 19-23, 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; Educational Change; *Educational Environment; Elementary Education; Longitudinal Studies; Models; *Reading Achievement; *Reading Improvement; Reading Research; Sex Differences; Teaching Methods
IDENTIFIERS *Accelerated Schools; Louisiana Educational Assessment Program

ABSTRACT

This study investigated improvement in reading/language arts test performance by non-transient learners in three accelerated schools. The sample included sixth-grade students who remained in the same accelerated school for at least four years. The third and fifth grade readers' scores on Louisiana's criterion-referenced test were used in this investigation. Logistic regression was used to explore the relationship of several variables to improvement in reading. A three-step sequential analysis was used to assess how various factors influenced reading improvement (i.e., whether students had a higher percentile ranking in sixth grade than in third grade). The first step included individual background variables (i.e., gender, retention, age, and base test scores). The second added the school environment variables. The third added the curriculum and instruction variables. Each step improved the fit of the model. Further, the differences between schools evident in step two were no longer significant in the third step, indicating curriculum and instructional practices explained the effects of school environments. Contains 24 references and 6 tables of data. (Author/RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 430 220

A Practical Model for Measuring the Effect of School Reform on the Reading Achievement of Non-Transient Learners

Presented at the Annual Meeting of the
American Educational Research Association

Montreal, Canada

April 22, 1999

by

Jill Berlin Slack
Southwest Educational Development Laboratory
Southeast Comprehensive Assistance Center

and

Edward P. St. John
Indiana Education Policy Center
Indiana University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. B. Slack

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

CS 013 582

Abstract

This study investigated improvement in reading/language arts test performance by non-transient learners in three accelerated schools. The sample included sixth grade students who remained in the same accelerated school for at least four years. The third and fifth grade readers' scores on Louisiana's criterion-referenced test were used in this investigation. Logistic regression was used to explore the relationship of several variables to improvement in reading. A three-step sequential analysis was used to assess how various factors influenced reading improvement (i.e., whether students had a higher percentile ranking in sixth grade than in third grade). The first step included individual background variables (i.e., gender, retention, age, and base test scores). The second added the school environment variables. The third added the curriculum and instruction variables. Each step improved the fit of the model. Further, the differences between schools evident in step two were no longer significant in the third step, indicating curriculum and instructional practices explained the effects of school environments.

Introduction

Several research reports have been written in the past decade demonstrating the importance of disaggregated test score analysis (Bamburg & Medina, 1993; GAO, 1991; McCarthy & Still, 1993; Murphy & Schiller, 1992). To date, however, most of this research has primarily concentrated on individual background characteristics such as race, gender, and socioeconomic class. Few studies have examined the influence of specific school practices on test score improvement. Such research can contribute valuable information for schools about the impact of their education practices, as well as for state and federal agencies concerned about reading improvement.

This paper reports a longitudinal analysis of the effects of a school reform on improvement in language arts achievement test data for Louisiana students who were engaged in the Accelerated Schools Project. The goals of the study were to assess: 1) the impact of specific school reform practices on student achievement, and 2) whether differences existed

between the test performances of students in three accelerated elementary schools.

This study pilot tests an approach to assessing the effects of site-based practices that can be used in schools to analyze the effects of locally designed school reforms. As background, we review related research. Then we describe the research method and study findings.

Background

In the early 1980s a concerted effort to reform American education began. The belief that the United States was falling behind other industrial countries in development, productivity, and quality was a dominant and recurring theme of the various education reform reports and proposals of the period (Carnegie Forum on Education and the Economy 1986; Education Commission of the States 1983; National Commission on Excellence in Education, 1983; National Governors' Association, 1986).

Focusing on the manner in which the American educational system competitively compared with that of the rest of the world, the National Commission on

Excellence in Education (1983) labeled the United States "*A Nation at Risk…*" and pointed to declines in student scores on standardized tests in its call for higher standards and educational reform to reverse the rising tide of mediocrity. As a result, the last decade has witnessed the proliferation of broad-ranged attempts to design forms of education that represent substantial departures from conventional practice: the Coalition of Essential Schools Program (Sizer, 1989); the Accelerated Schools Project (Hopfenberg et al., 1993; Levin, 1987); and the School Development Program (Comer, 1988), to name a few.

With pressure emanating from the national to the local level to improve student achievement, educators supporting and implementing reform programs have increasingly turned to standardized tests to prove the quality of their reform efforts. For example, Meister (1991) and McCarthy and Still (1993) used test scores to monitor the progress of accelerated schools. In addition, a General Accounting Office Report (GAO, 1991) used standardized test results to measure the effect of educational reform on high school students.

According to Herman and Winters (1991), "Standardized tests have been seen as the key measure of educational quality, satisfying the information needs of legislators and administrators at the federal, state, and local levels who wished to know how special programs were working and how schools were achieving" (p. 5). Consequently, "Tests have become so pervasive and the reform movement has invested so much importance in them, that increasingly it is in terms of standardized test scores alone that the nation judges its schools" (Toch, 1991, p. 206).

A concern flowing from this type of accountability is that test scores, by themselves, do not tell the whole story. Unless supplemented with other contextual factors, they have been found to provide a narrow account of school progress (Schmoker, 1996). Moreover, they have been found to have a limited influence on instruction (Haladaya et al., 1991; Smith, 1991).

In the 1990s research has placed more emphasis on disaggregated data analysis to find out how specific groups of students are impacted by various reform programs. However, these studies have primarily

focused on differences between races, ages (GAO, 1991), and socioeconomic classes (McCarthy and Still, 1993); and have produced mixed results. For example, the former study revealed no differences between race or age of students who were enrolled in schools with reform programs, whereas the latter study reported exceptional gains by low-socioeconomic students who attended schools with a reform program.

Hymes (1991) emphasized the importance of using disaggregated test data in the context of additional information. He found that responsible score interpretation takes into account demographic and process variables that help explain achievement disparities. Further, adequate attention to student and school characteristics increases the power of test scores to summarize and simplify, thus making them invaluable for analysis and discussion of educational reform (Peterson, 1992).

Bamburg and Medina (1993) also studied the impact of disaggregated test score analysis, but with an emphasis on teacher inquiry. They found the consistent use of disaggregated analysis increased the congruence between what teachers espoused and

practiced; focused the instructional conversation; and established data-driven priorities.

As indicated throughout this discussion, data analysis has many possibilities for helping schools reveal progress and areas that need improvement. Moreover, the impact of school reform depends on accurate methods of analyzing achievement test data. The traditional orientation, which involves characterizing schools' academic achievements based on aggregated test data does not consider many dynamics of the educational process, including school variation and school change.

This study extends the research on longitudinal analyses of disaggregated data by suggesting that non-transience and certain school variables, including school environment and instructional practices, influence student achievement.

Methodology

The purpose of the present study was the longitudinal investigation of language arts achievement test data of non-transient students and their schools in order to determine whether schools

with distinct contextual features experienced significantly different test score performances. The study involved the longitudinal analysis of non-transient sixth grade students, and included their third and fifth grade language arts raw scores from the Louisiana Educational Assessment Program (LEAP). Data were also collected from the second through fifth grade teachers on the instructional approaches they used on a daily basis with the students. This information was gathered through a survey and used in combination with school profile data to assist in the investigation.

Logistic regression was used to compare the test performance of three schools involved in this study. Only the non-transient students' test scores were considered in this investigation. Other school variables, such as individual background data and teachers' instructional approaches were also considered.

School Profiles

To assist in this investigation, profiles of the participating schools were developed using demographic

as well as process-oriented data relative to the Accelerated Schools Project. In addition, teachers completed a survey revealing the instructional approaches they used with the subjects in the study.

Three accelerated schools volunteered to participate in this study. The three schools are located in the Southeastern region of the United States and serve elementary students with similar socioeconomic status (SES) backgrounds. The schools are identified as X, Y, and Z.

Sixty-two, sixth grade, regular education students from these three schools were selected as subjects for the study because they were classified as "non-transient," that is, they remained in the same accelerated school for at least a four year period. Fourteen of 27 sixth graders in School X, 37 of 59 sixth graders in School Y, and 11 of 72 sixth graders in School Z qualified as non-transient students and therefore, subjects for the study. Table 1 provides specific details related to each school's subjects.

The dependent variable for this study was the non-transient student's language arts test scores on the LEAP. The independent variables focused on

individual background, school environment (the school), and the number of years subjects received each of five instructional approaches. The individual background variables included age, gender, retention information, and base test scores. Initially, race was considered. However, since 96% of the subjects were Black, this variable was omitted.

The instructional approaches included whole language, thematic units/teaching, multicultural education, tutoring, and Title I. The number of years subjects were engaged in a particular approach ranged between zero and four. Notably, Title I was a pull out program for the first three of the four years included in this study. In the last year, Title I became a schoolwide program (i.e., students received Title I services in their regular classrooms). Nonetheless, the Title I program provided subjects with additional reading instruction for all four years.

Table 1

Specific Details Relative to Each School's Participants

	School		
Participants	X	Y	Z
Number	14	37	11
% Black	100	100	82
% White			09
% Hispanic			09
% Female	50	62	82
% Retained	71	22	00
% Title 1	28	22	09

Statistical Methods

The dependent (outcome) variable was dichotomous with "0" representing no test score growth and "1" representing at least one point of test score growth between the third and fifth grade LEAP tests. One point was selected as the threshold for improvement to account for those subjects with high base scores and less room for improvement.

Most of the independent variables were also dichotomous either directly or through design sets.

For example, the number of years subjects were taught using the whole language approach represented a continuous variable (directly) that was grouped into three levels of dichotomous variables (a design set): two years, three years, and four years. The subjects either fell into a specific category (which was coded as a "1" or did not (which was coded as a "0").

Because the dependent variable in this analysis involved non-continuous outcome measures, a linear regression model, also termed Ordinary Least Squares regressions (OLS), was not considered an appropriate statistical procedure (Aldrich & Nelson, 1984). Instead, logistic regression, a form of multiple regression, was used to examine the relationship of specific variables to outcome measures (i.e., test score improvement).

Both the dependent variable and independent variables were used to develop the logistic regression model. The model used a sequential analysis, the stepping in of factors, to examine the relationship of specific factors to language arts test score improvement. The first step included only the individual background variables. The second step

added the school environment variables. The third step added the instructional variables.

The logic of each factor (individual background, school, and instructional approaches) was examined using two types of statistics. First, the changes in the delta Ps (probability measures) and the significance levels of each independent variable were reviewed. Second, the changes in the model statistics, particularly the log likelihood function (-2 Log L), pseudo R^2, and percent correctly predicted were compared across the model. The delta Ps were calculated using a formula recommended by Peterson (1984): $F(P)=\exp(L_1)/[1+\exp(L_1)]-\exp(L_0)/[1+\exp(L_0)]$ where L_0 = ln p/(1-p) (p = baseline probability, ln = natural logarithm) and $L_1 = L_0 + \text{Beta}$.

The delta P statistics were used in two ways in this investigation. First, for the dichotomous variables, the delta P provided a measure of probability on which a specific independent variable was likely to change the dependent variable. For example, if a specific characteristic had a delta P of 0.075, this was interpreted as increasing the probability of improvement by 7.5 percent for the

subjects having that characteristic. Likewise, a delta P of -0.075 was interpreted as decreasing the probability of improvement by 7.5 percent for the subjects with that characteristic.

The second use of the delta P statistic in this analysis was for continuous variables. In these cases, the delta P was interpreted as meaning that a change in unit measure altered the probability of the outcome by a specific percentage. For example, a delta P statistic of 0.095 for age was interpreted to mean that for each year of increase in age, the probability of improvement was 9.5 percent. A delta P statistic was calculated for each independent variable included in the logistic regression model.

In addition to the delta Ps, a pseudo R^2 and other model statistics such as the log likelihood function (-2 Log L) and the percent correctly predicted were applied to determine whether the model improved as variables were stepped in. Except for the pseudo R^2, the model statistics were provided on the Statistical Package for Social Sciences (SPSS) printout.

The pseudo R^2 was calculated using a method described by St. John, Kirshstein, & Noell (1991). Accordingly, the pseudo R^2 was computed by taking the -2 log likelihood statistic from the model and subtracting it from the -2 log likelihood statistic for the model containing the intercept only. This value was then divided by the -2 log likelihood statistic for the model containing the intercept only. An increase in the pseudo R^2 from one series over the last was interpreted to mean a reduction in unexplained error. Further, an increase in the percent correctly predicted and a decrease in the -2 Log L from one step to the next was interpreted as meaning an improvement in the overall predictability and fit of the model, respectively.

Model Specifications

Table 2 provides a listing of the specific coding for each variable. Specific independent variables are described in more detail below.

Individual Background: This factor included the following variables: age, gender, grade retention "once", grade retention "twice", and base test scores.

The numeric variable age represented age at the time of the last test administration.

The retention variables were determined through a design set of dichotomous variables due to their curvilinear relationship with test score improvement. Subjects were classified into one of three levels according to how many times they were retained.

School Environment: This category represented a design set of dichotomous variables that delineated whether the subjects attended School X, School Y, or School Z. More than half of the subjects in this study were from School Y. For this reason, subjects in School Y were uncoded (0/0) and compared to subjects in School X (1/0) and School Z (0/1).

Table 2

Variable Coding

Independent Variables		
Factor/Variable	**Coding**	**Comment**
Individual Background		
Age	Continuous	Indicates Age
Gender	0 = Male 1 = Female	Compares females to males
Retention "once"	0 = No 1 = Yes	Compares students who were retained once with students who were never retained
Retention "twice"	0 = No 1 = Yes	Compares students who were retained twice with students who were never retained
Base Scores	Continuous	Indicates Grade 3 Leap scores
School Environment		
School X	0 = No 1 = Yes	Compares School X with School Y
School Z	0 = No 1 = Yes	Compares School Z with School Y
Instructional Approach		
Whole Language (2 years)	0 = No 1 = Yes	Compares 2 years with 4 years
Whole Language (3 years)	0 = No 1 = Yes	Compares 3 years with 4 years
Thematic Units/ Teaching (3 years)	0 = No 1 = Yes	Compares 3 years with 1 year
Thematic Units/ Teaching (4 years)	0 = No 1 = Yes	Compares 4 years with 1 year

(table continues)

Multicultural Education (2 years)	0 = No 1 = Yes	Compares 2 years with 4 years
Multicultural Education (3 years)	0 = No 1 = Yes	Compares 3 years with 4 years
Tutoring (3 years)	0 = No 1 = Yes	Compares 3 years with 0 years
Tutoring (4 years)	0 = No 1 = Yes	Compares 4 years with 0 years
Title I (1 year)	0 = No 1 = Yes	Compares 1 years with 0 years
Title I (2 years)	0 = No 1 = Yes	Compares 2 years with 0 years
Title I (3 years)	0 = No 1 = Yes	Compares 3 years with 0 years

Instructional Approaches: The variables related to the instructional approaches included several design sets that were determined by the number of years subjects engaged in each of five approaches: whole language, thematic units/teaching, multicultural education, tutoring, and Title I. The number of years subjects were engaged in a particular approach ranged between zero and four years.

Except for Title I, the majority of the subjects in School Y received four years of each approach and the majority of the subjects in Schools X and Z received less than four years. Thus, the design sets were configured as such to create a balance and to prevent redundancies in the design matrix.

Findings

This longitudinal study investigated the achievement test performance of non-transient students in order to determine whether differences existed in the test performances of three distinct accelerated schools. Logistic regression was used to explore the relationship of several variables to achievement test performance. The variables included individual

background characteristics, school environment, and instructional approaches. Table 3 presents descriptive statistics for the variables by school. The statistics reveal the distinctiveness of the schools, especially with the teaching approaches.

A logistic regression model was developed to enhance the study. The dependent variables (i.e., LEAP scores) were coded as dichotomous outcomes, with "0" = no improvement and "1" = at least one point of improvement. One point was selected as the threshold to allow for subjects with high base test scores and smaller margins for improvement.

Table 4 presents the population with LEAP Language Arts test score improvement broken down by the amount of improvement. As the table reveals, of the participants who experienced improvement, only 3% fell into the "+1" point category and 67% fell into the "$\geq$5" points category. Moreover, most of the subjects with the lower base scores (representing the bottom half) fell into the "$\geq$5" category and most with higher base scores (representing the top half) fell into the "+1" to "+4" categories. This suggest that a

Table 3

Variables by School

School (Environment)	X	Y	Z
Individual Background			
Age			
10	21%	38%	45%
11	36%	62%	55%
12	36%		
13	7%		
Gender			
Female	50%	62%	82%
Retention "Once"	50%	22%	
Retention "Twice"	21%		
Base Test Score			
LEAP LA Grade 3	353.0	357.0	357.0
Instructional Approaches			
Whole Language			
2 years	100%		73%
3 years		8%	27%
4 years		92%	
Thematic Units/Teaching			
1 year	100%		
2 years		8%	91%
3 years		92%	9%
Multicultural Education			
2 years			82%
3 years	100%	8%	18%
4 years		92%	
Tutoring			
0 years			100%
3 years	100%	5%	
4 years		95%	
Title I			
0 years	72%	76%	91%
1 years		16%	
2 years	14%	8%	9%
3 years	14%		

Table 4

Population with LEAP Language Arts Test Score Improvement Broken Down by Amount of Improvement

Point Category	Percent
+1	3.0
+2	3.0
+3	13.5
+4	13.5
≥5	67.0

one point threshold for improvement (rather than a two or more point threshold) may have strengthened the design of the logistic regression model, allowing for a better comparison between the growth of subjects with high base scores and those with low base scores.

Like the dependent variables, most of the independent variables were also coded as dichotomous variables, with 1 = yes, and 0 = no. Of the individual background variables, females were compared to males, and students who were retained one or two years were compared with students who were never

retained. Age and base test score were continuous variables.

Due to the unbalanced number of subjects representing the schools (14 in School X, 37 in School Y, and 11 in School Z), the logistic regression model was designed to compare School X and Z with School Y.

Additionally, the last step of the model was designed to compare each group who received less than four years of whole language and multicultural education with the group who received four years. To counteract redundancies in the design matrix, the groups who received three and four years of thematic units/teaching and tutoring were compared with the group who received the fewest years. Finally, each group who received Title I services was compared with those who did not receive such services.

The factors of individual background, school environment, and instructional approaches were added sequentially to the logistic regression model to examine their additive effects on achievement test performance. The logic of each factor was examined using two types of statistics.

First, the changes in the delta Ps and the significance levels of each independent variable (associated with improvement) were reviewed. Second, the changes in the model statistics, particularly the log likelihood function (-2 Log L), pseudo R^2, and percent correctly predicted were compared across the series.

Table 5 presents the results of the sequential analysis for the LEAP Language Arts Model. In step one, the LEAP base scores variable was significant at the 0.05 level ($p \leq .05$). The students who had higher language arts scores in third grade were 2.1 percent less likely to improve than the students who had lower language arts scores.

In step two, the language arts base scores remained significant and resulted in a 2.6 percent decrease in the probability of improvement for the students with higher scores ($p \leq .01$). One of the school environment variables also had a significant and negative association with improvement. Students in School X were 37.6 percent less likely to improve than the students in School Y ($p \leq .05$).

Table 5

Logistic Analysis of Test Score Improvement - LEAP Language Arts Model

Factor/Variable	Step 1 Delta P	Step 2 Delta P	Step 3 Delta P
Individual Background			
Age	0.130	0.169	0.276*
Female	-0.107	-0.133	-0.093
Retention			
1 Year	-0.192	-0.024	-0.105
2 Years	-0.302	-0.012	0.439
LEAP 3 LA Scores	-0.021**	-0.026***	-0.030**
School Environment			
School X		-0.376**	0.440
School Z		-0.094	0.440
Instructional Approaches			
Whole Language			
2 years			-0.301
3 years			0.440
Thematic Units/Teaching			
3 years			-0.440
4 years			0.217
Multicultural Education			
2 years			0.329
3 years			0.437
Tutoring			
3 years			-0.440
4 years			0.440
Title I			
1 year			-0.352*
2 years			-0.217
3 years			-0.440
-2 Log	73.113	65.750	50.801
Pseudo R^2	0.097	0.188	0.373
% Correctly Predicted	64.41	76.27	79.66
Model Chi Square	7.846	15.209	30.157
Goodness of Fit	57.347	64.215	62.843

* = 0.10 level of significance
** = 0.05 level of significance
*** = 0.01 level of significance

The -2 Log L for this step decreased from the previous step, indicating a more appropriate fit. Additionally, the pseudo R^2 increased from 0.097 to 0.188 and the percent correctly predicted improved from 64.41 to 76.27.

When the instructional variables were added in step three, the base scores variable continued to be significant and negative. The students who had higher scores in third grade were 3.0 percent less likely to improve than those who had lower scores ($p \leq .05$).

One of the most interesting changes in this step was the significance of age. Students were 27.6 percent more likely to improve for each additional year of age ($p \leq .10$). This finding suggests an interaction between the students' ages and the instructional approach variables.

Another interesting observation is the significant, negative association related to Title I. Students who received one year of Title I services were 35.2 percent less likely to improve than the students who did not receive Title I services ($p \leq .10$).

The change in the -2 Log L indicates that step three of the model provided the best fit. The

increase in the pseudo R^2 to 0.373 and the percent correctly predicted to 79.66 also suggests that step three of the model was significantly better at predicting improvement than the previous two steps.

Conclusion

This study extends the research on longitudinal analysis of disaggregated data by developing a logistic regression model to investigate whether specific variables related to school reform impact the test performance of non-transient learners. In examining the variables across the regression model used in this study, several points emerged.

First, there is a consistent, negative association between the base test score variable and improvement. Students who had high scores on the third grade LEAP were less likely to improve than students who had low scores. This finding suggests that the Accelerated Schools process may work especially well for low-performing students.

Table 6 provides further evidence of the benefits of accelerating the learning of low-performing students, that is, students who scored at or below 350

(i.e., the bottom third) on the third grade LEAP. The table presents the amount of points these students improved from the third to the fifth grade on the LEAP Language Arts test. As the table reveals, 73 percent of the low-performing students improved their scores by five points or more and 64% of them improved their scores by ten points or more.

A second point that emerged from the study relates to the school environment factor. In step two, there was a negative association between

Table 6

Percent of Low-Performing Students Experiencing Improvement

Points Improved	LEAP Language Arts
≥5	72.7
≥10	63.7

Note. Low performing refers to those students who scored at or below 350 on the third grade LEAP Language Arts section.

improvement and School X. Students in School X were 38 percent less likely to improve than students in School Y.

However, the differences between the schools were no longer significant when the instructional approach factor was considered. This finding demonstrates an interaction between student engagement in the instructional approaches and their school environment. That is, student engagement in the instructional approaches may compensate for their school environment. This further suggests that what actually goes on in schools, besides matters related to curriculum and instruction, influences test performance.

A third point that emerged is the positive association between age and test score improvement. In particular, older students were more likely to improve. The probability of improvement increased 28 percentage points for each year of age. Aside from the retention variables, this suggests that students may perform better academically (as measured by achievement tests) when they start school at an older age than the norm age, say for kindergarten.

Interestingly, age only became significant when the instructional approach factor was added to the model. This change demonstrates a relationship between students' ages and the instructional variables. Since all of the older (12 and 13 year old students) attended School X, their engagement in the instructional approaches may have influenced their improvement in language arts.

Additionally, one year of Title I was found to have a significant, negative influence on test score performance. That is, students who received Title I services for one year were 35 percent less likely to improve than students who never received Title I services.

These findings do not suggest that the other instructional approaches in this study were not effective at improving test scores. Rather they suggest that one year to three years differential in the amount of exposure to the instructional approaches may not have a significant impact on improvement.

A final point relates to the model statistics. The substantial decrease in the -2 Log L statistic suggests that the inclusion of more variables improved

the fit of the model. Furthermore, the steady increases in the pseudo R^2 and percent correctly predicted indicate that the model became significantly better at predicting improvement with each additional step.

The logistic regression model proposed in this study is both practical and easy to use. Further, it offers teachers, administrators, and policymakers insight on creative ways of measuring schools progress. The model also allows educators to determine whether, and to what degree, school environment and instructional variables have an impact on student achievement. This innovative orientation of measuring school progress can provide a crucial missing link in helping schools, especially those engaged in reform, learn more about themselves.

The availability of workable models for school-based research on reading improvement may be increasingly important in the early 21st century, given the new federal Reading Excellence Act and the emergence of state-funded reading improvement programs (e.g., St. John, Bardzell, Michael, Hall, Manoil, Asker, & Clements, 1998). Increasingly, schools must

propose "research-based" programs to secure funding. The model pilot tested in this paper can be used by schools interested in developing their own research-based approach.

References

Aldrich, J. H. & Nelson, F. D. (1984). Linear probability, logit and probit models. Beverly Hills, CA: Sage Publications.

Bamburg, J. & Medina E. (1993). Analyzing student achievement: Using standardized tests as the first step. In J. Bamburg (Ed.), Assessment: How do we know what they know? Dubuque, IA: Kendall-Hunt.

Carnegie Forum on Education and the Economy. (1986). A nation prepared: Teachers for the 21st century. Washington, DC: The Forum.

Comer, J. P. (1988). Educating poor minority children. Scientific American, 259(5), 42-48.

Education Commission of the States. (1983). Action for excellence. Denver, CO: The Commission.

General Accounting Office. (1991). The effect of education reform on student achievement: GAO report. Urban Education, 26, 160-175.

Haladyna, T. M, Nolen, S. B. & Haas, N. (1991). Raising standardized achievement test scores and the origins of test score pollution. Educational Researcher, 20(5), 2-7.

Herman J. L., & Winters, L. (1991). Sensible school-based evaluation: Multilevel evaluation systems project (Report No. TM-016-200). Washington, DC: Center for Research on Evaluation, Standards, and Student Testing. (ERIC Document Reproduction Service No. ED 329 575)

Hopfenberg, W. S., Levin, H. M., Chase, C., Christensen, G., Moore, M., Soler, P., Brunner, I., Keller, B., & Rodriguez, G. (1993). Accelerated schools resource guide. San Francisco: Jossey-Bass.

Hymes, D. L. (1991). The changing face of testing and assessment: Problems and solutions. Arlington, VA: American Association of School Administrators.

Levin, H. (1987). Accelerated schools for disadvantaged students. Educational Leadership, 44(6), 19-21.

McCarthy, J., & Still, S. (1993). Hollibrook accelerated elementary school. In Joseph Murphy & P. Hallinger (Eds.), Restructuring schooling (pp. 63-83). Newbury Park, CA: Corwin Press.

Meister, G. (1991). Assessment in programs for

disadvantaged students: Lessons from accelerated schools (Report No. TM-018-032). Washington, DC: U.S. Congress, Office of Technology Assessment. (ERIC Document Reproduction Service No. ED 340 776)

Murphy, J., & Schiller, J. (1992). Tools for diagnosis, prescription, and accountability. In J. Murphy & J. Schiller (Eds.), Transforming America's schools (pp. 187-224). LaSalle, IL: Open Court.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative of educational reform. Washington, DC: U.S. Government Printing Office.

National Governors' Association. (1986). Time for results. Washington, DC: The Association.

Peterson, T. (1984). A comment on presenting results of logit and probit models. American Sociological Review, 50(1), 130-131.

Peterson, T. K. (1992). Designing accountability to help reform. In C. E. Finn, Jr. & T. Rebarber (Eds.), Education reform in the '90s (pp. 109-132). New York: Macmillan.

St. John, E. P., Bardzell, J., Michael, R., Hall, G., Manoil, K., Asker, E., & Clements, M. (1998). Indiana's Early Literacy Intervention Grant Program. Bloomington, IN: Indiana Education Policy Center.

St. John, E. P., Kirshstein, R. J., & Noell, J. (1991). The effects of student financial aid on persistence: A sequential analysis of the high school and beyond senior cohort. Review of Higher Education, 14(3), 383-486.

Schmoker, M. (1996). Results: The key to continuous school improvement. Alexandria, VA: Association for Supervision and Curriculum Development.

Sizer, T. R. (1989). Diverse practice, shared ideas: The essential school. In H. Walberg & J. Lane (Eds.), Organizing for learning: Toward the 21st century. Reston, VA: National Association of Secondary School Principals.

Smith, M. (1991). Put to the test: The effects of external testing on teachers. Educational Researcher, 20(5), 8-11.

Toch, T. (1991). The coins of realm: The paradox of standardized testing. In T. Toch (Ed.), In the

name of excellence (pp. 205-232). New York: Oxford University Press.